

# Weekly Report 2016.05.22-2017.05.30

## Progress

### 1. 大图可视化调研

本周一与嘉铨，东明，于晨进行了讨论，我们认为可以分头去对大图可视化中不同的技术领域进行调研。主要包括三个方面：首先，图数据库调研；其次，直接查询，局部可视化的图可视化工具调研；最后，整体进行可视化，层次细节展示的图可视化工具调研。我们进行了分工，并分头进行调研工作。

### 2. TCP Tree 文章修改

本周基本完成了对 TCP Tree 文章的修改。

对 TCP Tree 应用到其他数据上的一些想法：

我们目前使用的空气污染数据以及计算出来的传感器/属性间相互影响的数据实际上可以被抽象为一种比较特殊的动态图数据：即节点的属性不断变化（空气污染属性的变化）以及边属性的不断变化（相互影响关系）。这一点我们最初的可视设计也有体现，那个设计实际上就是一个全联通图加上边粗细的编码。

TCP Tree 提供了一种数据组织的方式，可以把动态网络数据的拓扑结构和时序变化统一到一个结构中去，然后提供不同的分割方式让用户在拓扑结构上和时序上对网络进行探索。我在文章中写的是两种操作：correlation partition 和 time partition。如果将目前的工作推广到动态图上的话，这两种操作实际上对应的是两个不同的操作集合，每个集合中可能都包含着一系列不同的分割方法，每种分割方法可能对应了不同的分割对象。

对于动态网络来说，其节点属性，边属性，以及拓扑结构都可以随时间变化。原来的 correlation partition 推广到动态网络下就是对图拓扑结构的切割，比如这次 vis 投稿中的树切割算法；而 time partition 则是针对时序变化中，具有某些属性的特定序列的切割。

综上，TCP Tree 应该能够推广到动态网络上。

同时，我也跟琦神和天野讨论了将 TCP Tree 应用到电网数据的可能性。天野提到了需要对电网数据中的异常进行可视分析。我觉得这种异常检测其实也可以视为 TCP Tree 中 time partition 的一种方法，切出来的是可能包含异常的时间片段

### 3. 图/动态图数据可视化构建的想法

这几次 vis 投稿的过程中，我有一个感触，就是质量比较好的图/动态图数据是非常难找的。现有的一些开放数据要么就是数据量太大，要么就是结构不太理想，导致我做来做去总是在基于 DBLP 数据集在做。同时，没有数据也会影响到我们开发系统时的进度。

在这次 vis 投稿中，我们在做 user study 时自己构造了一些测试用的假数据，来降低我们在 user study 时的任务难度。沿着这个思路，我觉得我们也许可以做一个小的研究题目/系统，专门来做图/动态图的数据生成。